

Correntropy-based Nonlinearity Test Applied to Patients with Chronic Heart Failure

Ainara Garde, *Student Member, IEEE*, Leif Sörnmo, *Senior Member, IEEE*, Raimon Jané, *Member, IEEE*,
and Beatriz F. Giraldo, *Member, IEEE*

Abstract—In this study we propose the correntropy function as a discriminative measure for detecting nonlinearities in the respiratory pattern of chronic heart failure (CHF) patients with periodic or nonperiodic breathing pattern (PB or nPB, respectively). The complexity seems to be reduced in CHF patients with higher risk level. Correntropy reflects information on both, statistical distribution and temporal structure of the underlying dataset. It is a suitable measure due to its capability to preserve nonlinear information. The null hypothesis considered is that the analyzed data is generated by a Gaussian linear stochastic process. Correntropy is used in a statistical test to reject the null hypothesis through surrogate data methods. Various parameters, derived from the correntropy and correntropy spectral density (CSD) to characterize the respiratory pattern, presented no significant differences when extracted from the iteratively refined amplitude adjusted Fourier transform (IAAFT) surrogate data. The ratio between the powers in the modulation and respiratory frequency bands R was significantly different in nPB patients, but not in PB patients, which reflects a higher presence of nonlinearities in nPB patients than in PB patients.

I. INTRODUCTION

Chronic heart failure (CHF) is often related to breathing abnormalities that, according to several studies, may reflect the condition of the patients with CHF [1]. The respiratory patterns developed by CHF patients include various forms of oscillatory breathing patterns characterized by rises and falls in ventilation [2]. Previous studies present periodic breathing (PB) during sleep or wakefulness as a powerful predictor of poor prognosis in patients with CHF [3]. Periodic breathing patterns can be classified into ventilation with apnea, commonly known as Cheyne Stokes respiration (CSR), or ventilation without apnea [4], [5]. Various studies report a PB prevalence as high as 70% in CHF patients [6]. Moreover PB, and especially CSR, have been associated with increased mortality [1], [7]. Accurate risk stratification establishing patient prognosis is crucial for appropriate allocation of limited resources. Physiological parameters for the

characterization and detection of different breathing patterns have been suggested in a number of clinical studies [8].

Evidence for nonlinearities in cardiac physiology includes the finding of abrupt changes, self-sustained and complex oscillations, and chaotic behavior. Nonlinear indices have improved the assessment of complex autonomic coordinations in several cases as risk stratification of CHF patients. According to several studies a nonlinear approach can improve early risk identification of cardiovascular events such as sudden cardiac death. Different pathological situations seem to have different complexity and regularity properties. Complexity seems to be reduced in CHF patients with a higher risk condition [9], [10], [11] and [12].

Normal breathing cycle lengths range from 3 to 5 s (i.e., 0.20–0.33 Hz) and PB patterns cycle lengths from 25 to 100 s (i.e., 0.01–0.04 Hz) [13]. Both the respiratory frequency and modulation frequency appear to be essential for the understanding of periodic and nonperiodic breathing (nPB) patterns in CHF patients.

Recently introduced by Santamaria et al. [14], correntropy is a similarity measure defined in terms of inner products of vectors in a kernel parameter space. It involves a positive definite kernel function which implicitly transforms the original signal into a high dimensional space in a nonlinear way and calculates efficiently a generalized correlation in that space. Correntropy combines the statistical distribution and the time structure of the signals into a single measure. Unlike autocorrelation, it can preserve higher order moments and therefore the existing nonlinearities, which makes correntropy a well-qualified candidate for characterization of nonlinear dynamics.

Surrogate data analysis is a popular tool, usually exploited to evaluate the existence of nonlinear dynamics. The procedure is based on a null hypothesis to be rejected and the construction of a surrogate data test according to the null hypothesis. To reject the null hypothesis, a statistical test is applied to a discriminating parameter calculated from the original and surrogate data [15].

In this study we propose correntropy as a discriminant measure to study the nonlinear structure in CHF patients with periodic breathing and nonperiodic breathing pattern through the respiratory flow signal, motivated by its ability to preserve nonlinear characteristics and computational simplicity in high dimensional spaces.

This work was supported in part by Ministerio de Ciencia e Innovación under grants TEC2007-68076-C02-01 and TEC2007-63637 from the Spanish Government.

A. Garde, B.F. Giraldo and R. Jané are with Dept. of ESAT, Universitat Politècnica de Catalunya (UPC), Institut de Bioenginyeria de Catalunya (IBEC) and CIBER de Bioingeniería, Biomateriales y Nanomedicina (CIBER-BBN). c/. Pau Gargallo, 5, 08028, Barcelona, Spain. (fax: +34 93 401 7045) e-mail: ainara.garde@upc.edu, beatriz.giraldo@upc.edu, raimon.jane@upc.edu

L. Sörnmo is with the Department of Electrical and Information Technology and Center of Integrative Electrophysiology, Lund University, Lund SE-221 00, Sweden (e-mail: leif.sornmo@es.lth.se).

II. METHODS

A. Correntropy

For a discrete-time, strictly stationary stochastic process $x(n)$, $n = 0, \dots, N-1$, where N is the number of samples, the correntropy function $V(m)$ is defined as [14], [16]:

$$V(m) = E[\kappa(x(n), x(n-m))], \quad (1)$$

where $E[\cdot]$ is the statistical expectation operator and $\kappa(\cdot)$ is a symmetric, positive definite kernel function. In this study, the centered correntropy function $V_c(m)$ is employed in which the mean of the transformed data is subtracted so as to reduce the effect of output DC bias [16]. Its definition is given by

$$V_c(m) = E_{x(n)x(n-m)}[\kappa(x(n), x(n-m))] - \quad (2)$$

$$E_{x(n)}E_{x(n-m)}[\kappa(x(n), x(n-m))], \quad (3)$$

where $E_{x(n)}E_{x(n-m)}[\kappa(x(n), x(n-m))]$ is the square mean of the transformed data, corresponding to the correntropy mean value (\bar{V}) which can be estimated through the kernel function from the original data,

$$\bar{V} = \frac{1}{N^2} \sum_{n=1}^N \sum_{m=n+1}^N \kappa(x(n), x(n-m)). \quad (4)$$

The Gaussian kernel function, applied in the present study, is given by

$$\kappa(x(n), x(n-m)) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{(x(n) - x(n-m))^2}{2\sigma^2}\right], \quad (5)$$

where σ is referred to as the kernel parameter (“standard deviation”). With the Taylor series expansion of the Gaussian kernel, we can rewrite the correntropy function as

$$V(m) = \frac{1}{\sqrt{2\pi}\sigma} \sum_{k=0}^{\infty} \frac{(-1)^k}{2\sigma^{2k}k!} E[(x(n) - x(n-m))^{2k}], \quad (6)$$

containing all even-order moments of the random variable $(x(n) - x(n-m))$.

Different kernel functions would yield different expansions, but most kernel functions are nonlinear and therefore include higher-order statistical information about the stochastic process. From (6) it is obvious that the emphasis given to higher-order moments is controlled by σ . In this study, σ is determined by Silverman’s rule of density estimation [14]:

$$\sigma = 0.9AN^{-1/5}, \quad (7)$$

where A is the smaller value of the standard deviation of data samples and the data interquartile range scaled by 1.34, and N is the number of samples. Since the flow signals are scalar, this rule is simple to apply and is known to produce reasonable values. Hence, the sample estimator of $V(m)$ in (1) is given by

$$\hat{V}(m) = \frac{1}{N-m+1} \sum_{n=m}^N \kappa(x(n), x(n-m)), \quad (8)$$

The centered correntropy estimated by $\hat{V}_c(m) = \hat{V}(m) - \bar{V}$ is therefore used throughout the study.

A powerful advantage of the correntropy is its robustness against impulsive noise which is due to that the Gaussian kernel makes $\kappa(x(n), x(n-m)) \approx 0$ when either $x(n)$ or $x(n-m)$ is an outlier. Thanks to this attractive property, the results of the present study were obtained without any kind of preprocessing. As the respiratory frequency can range up to 0.5 Hz, the respiratory flow signal is downsampled to 2 Hz.

B. Correntropy Spectral Density

Similar to the conventional correlation function, it can be shown that correntropy is a positive definite function and therefore lends itself for many signal processing applications. Since we are interested in exploring the spectral properties of the respiratory signals, such information is conveyed by the nonparametric correntropy spectral density (CSD), defined by:

$$\hat{P}_x(e^{j\omega}) = \sum_{m=-(N-1)}^{N-1} \hat{V}_c(m) e^{-j\omega m}. \quad (9)$$

Considering that the flow signal is oscillatory in nature with a spectrum that exhibits pronounced peaks, autoregressive (AR) spectral analysis is here applied as it offers better spectral resolution than does periodogram-based spectral analysis. The correntropy extension to a parametric CSD, assuming that the signal is characterized by an AR model, is straightforward because the correlation properties are preserved.

C. Parameter extraction

The respiratory pattern is characterized in the time domain by the correntropy mean, i.e., Renyi’s entropy and in frequency domain by power related parameters derived from the CSD. Two frequency bands are defined, being centered around the respiratory frequency (“ Δf_r : the respiratory frequency band”) and the modulation frequency (“ Δf_m : the modulation frequency band”). The modulation frequency of PB patterns ranges typically from 0.005 to 0.05 Hz, and the modulation frequency peak (f_{p_m}) is located in this band with $\Delta f_m = 0.02$ Hz. The respiratory frequency peak (f_{p_r}) is tracked in the band from 0.2 to 0.4 Hz with $\Delta f_r = 0.2$ Hz. The individual powers of the two frequency bands, P_m and P_r , and their ratio $R = P_m/P_r$ constitute the spectral parameters investigated in this study, see Table I.

TABLE I
PARAMETER DESCRIPTION

Parameter	Description
P_m	Power of the modulation frequency band
P_r	Power of the respiratory frequency band
R	Relation between the P_m/P_r
\bar{V}	Correntropy mean

D. Surrogate for nonlinearity test

The null hypothesis considered in this study is that the analyzed data is generated by a Gaussian linear stochastic process. Therefore, surrogates that keep both the power

spectrum and the amplitude distribution of the original data are the best suited in order to analyze the nonlinearity and complexity of the flow signals [15], [17]. These surrogate data are usually created to have the same Fourier amplitudes as the original data but with random phases. The conventional power spectral density is the squared amplitude of the Fourier transform; therefore the original and its surrogate data obtained by this method share the same power spectral densities, and therefore the same autocorrelation function. By such phase randomization, linear correlations are maintained, but any underlying nonlinear dynamic structure within the original data is altered [18].

Iteratively refined amplitude adjusted Fourier transform (IAAFT) algorithm for generating surrogates, maintains the linear correlations and the power spectral density of the original data, but eliminates the nonlinearities. The surrogates throughout this study are generated using this technique. The value of any nonlinear measure evaluated on the original and surrogate data is compared. If significant differences are shown between both data, the null hypothesis is rejected. For a level of significance of α , a set of $M = 1/\alpha - 1$ surrogate sequences need to be generated for a one-sided test [19]. For testing at a 95% level of significance ($\alpha = 0.05$), at least 39 surrogates are required. Engaging more surrogates increases the confidence level of a test. Mann-Whitney test is used to reject the null hypothesis using the parameters extracted from the respiratory flow signal and its surrogate data.

III. EXPERIMENTAL PROTOCOL AND DATA ANALYSIS

A. Respiratory datasets and pre-processing

Respiratory flow signals were recorded from 35 healthy volunteers (12 males, 23 females, aged 26.6 ± 7 years) and 26 patients with CHF (19 males, 7 females, aged 65.0 ± 9 years) at the Santa Creu i Sant Pau Hospital in Barcelona, Spain. All subjects were studied according to a protocol previously approved by the local ethics committee. The respiratory flow signal was acquired using a pneumotachograph, consisting of a Datex-Ohmeda monitor with a Validyne Model MP45-1-871 Variable-Reluctance Transducer. The signals were recorded at 250 Hz sampling rate.

According to clinical criteria, CHF patients were classified into two groups: 8 patients with periodic breathing pattern and 18 patients with nonperiodic breathing pattern.

B. Data analysis

We proposed the CSD as a tool to detect nonlinear structures into respiratory pattern. The IAAFT was used to generate the surrogate data. For testing at a 95% level of significance ($\alpha = 0.05$), at least 39 surrogates were required, but we chose to use 150 surrogates ($\alpha = 0.013$) in order to increase the discrimination power [19].

In order to characterize the respiratory pattern, the parameters $R = P_m/P_r$ and \bar{V} are derived from the correntropy and CSD. The nonlinearities reflected by these parameters are evaluated by a Mann-Whitney test comparing the parameters extracted from the original flow signal and its surrogate data.

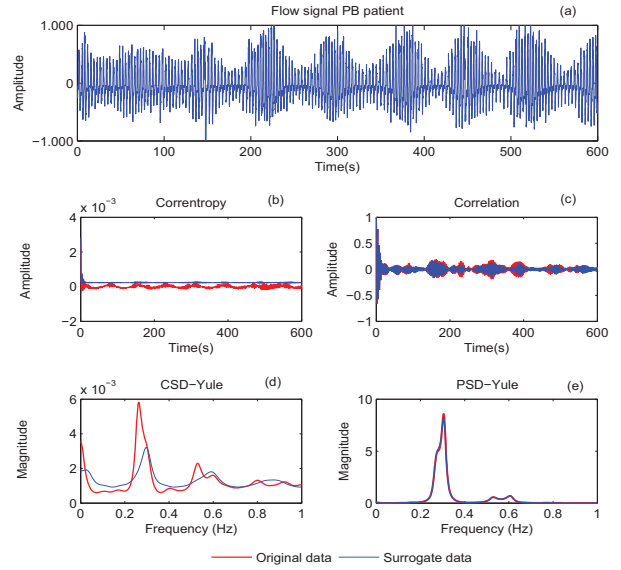


Fig. 1. (a) The respiratory flow signal of a CHF patient with PB, (b) correntropy, and (c) correlation, (d) the CSD and (e) the PSD, both obtained with Yule-Walker's method, computed for the original and surrogate data.

This nonlinearity presence is evaluated by classifying PB and nPB patterns in CHF patients.

IV. RESULTS

Figs. 1 and 2 illustrate the performance of the conventional correlation and PSD, and the correntropy and CSD calculated on the original respiratory flow signal and its corresponding surrogate data, of subjects with PB and nPB patterns, respectively. As expected, the autocorrelation and PSD, obtained with flow signals and their surrogate data, are approximately equal for all cases. However, significant differences are observed in both correntropy and CSD of the original and surrogate data which reflects the presence of nonlinearities in respiratory flow signal.

Table II shows the mean and the standard deviation of the parameters derived from correntropy and CSD, using the original flow signal from CHF patients with periodic and nonperiodic breathing pattern and its IAAFT surrogate data. The parameter R evaluated with the original and surrogate data present significant differences ($p = 0.003$) in patients with nPB patterns but not in patients with PB patterns.

V. DISCUSSION

This paper investigates the complexity of PB and nPB patterns in CHF patients in terms of nonlinearity through correntropy and CSD. In contrast to the conventional second-order correlation function, the correntropy function accounts for higher-order moments and reflects the nonlinearities of the data. Therefore, we propose correntropy as surrogate-based nonlinearity test. The respiratory pattern is characterized by some parameters derived from correntropy and CSD applied to respiratory flow signal.

In our recent studies [20] and [21], the ratio between the powers in the modulation and respiratory frequency bands R

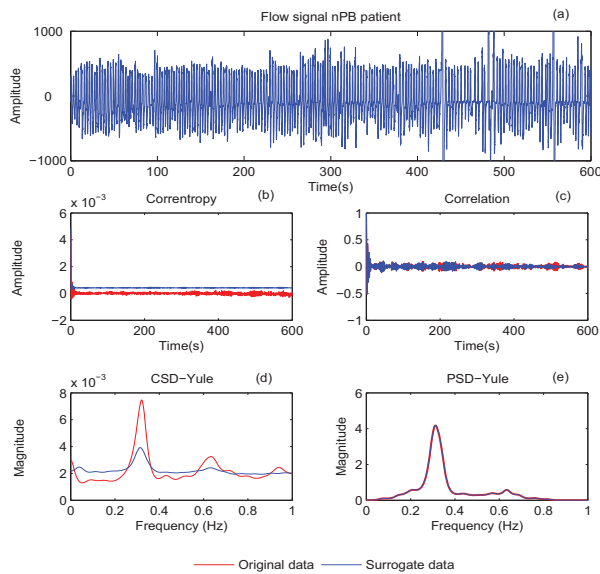


Fig. 2. (a) The respiratory flow signal of a CHF patient with nPB, (b) correntropy, and (c) correlation, (d) the CSD and (e) the PSD, both obtained with Yule-Walker's method, calculated over the the original and surrogate data.

TABLE II

MEAN \pm STANDARD DEVIATION AND NONLINEARITY PRESENCE (p -VALUE BETWEEN ORIGINAL AND SURROGATE DATA) OF THE PARAMETERS R AND \bar{V}^* CALCULATED ON THE RESPIRATORY FLOW OF CHF PATIENTS WITH PB AND nPB AND THEIR CORRESPONDING SURROGATE DATA.

	Original data	Surrogate data	p -value
PB			
R	0.14 ± 0.58	0.093 ± 0.012	n.s.
\bar{V}^*	0.3 ± 0.12	0.3 ± 0.12	n.s.
nPB			
R	0.071 ± 0.025	0.089 ± 0.012	0.0003
\bar{V}^*	0.4 ± 0.16	0.4 ± 0.16	n.s.

*: The values have been multiplied by 1000 for ease of legibility

provided an accuracy of 88.9% when classifying PB versus nPB. In this study, we have observed that the ratio R extracted from the original signal and the one extracted from its surrogate data present significant differences in nPB patients, but not in PB patients, thus reflecting a higher degree of nonlinearity in nPB patterns. The small size of the dataset is a limitation of the present study. The significance of the results, needs to be further established on a larger set.

VI. ACKNOWLEDGMENTS

The authors would like to thank to Drs. S. Benito and A. Bayés-Genis and their collaborators of Santa Creu i Sant Pau Hospital, Barcelona, Spain, for their collaboration in the signal database acquisition.

REFERENCES

[1] G. D. Pinna, R. Maestri, A. Mortara, and M. T. La Rovere, "Cardiorespiratory interactions during periodic breathing in awake chronic heart failure patients," *Am. J. Physiol. Heart Circ. Physiol.*, vol. 278, no. 3, pp. H932–41, 2000.

[2] T. Yajima, A. Koike, K. Sugimoto, Y. Miyahara, F. Marumo, and M. Hiroe, "Mechanism of periodic breathing in patients with cardiovascular disease," *Chest*, vol. 106, no. 1, pp. 142–146, 1994.

[3] P. A. Lanfranchi, A. Braghiroli, E. Bosimini, G. Mazzuero, R. Colombo, C. F. Donner, and P. Giannuzzi, "Prognostic value of nocturnal Cheyne-Stokes respiration in chronic heart failure," *Circulation*, vol. 99, no. 11, pp. 1435–1440, 1999.

[4] G. Lorenzi-Filho, P. R. Genta, A. C. Figueiredo, and D. Inoue, "Cheyne-stokes respiration in patients with congestive heart failure: causes and consequences," *Clinics (Sao Paulo, Brazil)*, vol. 60, no. 4, pp. 333–344, 2005.

[5] D. P. Francis, K. Willson, L. C. Davies, A. J. Coats, and M. Piepoli, "Quantitative general theory for periodic breathing in chronic heart failure and its clinical implications," *Circulation*, vol. 102, no. 18, pp. 2214–2221, 2000.

[6] G. D. Pinna, R. Maestri, A. Mortara, P. Johnson, T. Witkowski, P. Ponikowski, D. Andrews, S. Capomolla, M. T. La Rovere, and P. Sleight, "Nocturnal periodic breathing is an independent predictor of cardiac death and multiple hospital admissions in heart failure," in *Proc. Comput. Cardiol.*, vol. 33, 2006, pp. 837–840.

[7] P. J. Hanly and N. S. Zuberi-Khokhar, "Increased mortality associated with cheyne-stokes respiration in patients with congestive heart failure," *Am. J. Resp. Critical Care Medicine*, vol. 153, no. 1, pp. 272–276, 1996.

[8] J. P. Ribeiro, "Periodic breathing in heart failure: bridging the gap between the sleep laboratory and the exercise laboratory," *Circulation*, vol. 113, no. 1, pp. 9–10, 2006.

[9] K. K. L. Ho, G. B. Moody, C.-K. Peng, J. E. Mietus, M. G. Larson, D. Levy, and A. L. Goldberger, "Predicting Survival in Heart Failure Case and Control Subjects by Use of Fully Automated Methods for Deriving Nonlinear and Conventional Indices of Heart Rate Dynamics," *Circulation*, vol. 96, no. 3, pp. 842–848, 1997.

[10] M. A. Woo, W. G. Stevenson, D. K. Moser, R. B. Trelease, and R. M. Harper, "Patterns of beat-to-beat heart rate variability in advanced heart failure," *Am. Heart J.*, vol. 123, no. 3, pp. 704 – 710, 1992.

[11] S. Guzzetti, S. Mezzetti, R. Magatelli, A. Porta, G. D. Angelis, G. Rovelli, and A. Malliani, "Linear and non-linear 24 h heart rate variability in chronic heart failure," *Auton. Neurosci.*, vol. 86, no. 1-2, pp. 114–123, 2000.

[12] S. Guzzetti, R. Magatelli, E. Borroni, and S. Mezzetti, "Heart rate variability in chronic heart failure," *Auton. Neurosci.*, vol. 90, no. 1-2, pp. 102 – 105, 2001.

[13] G. D. Pinna, R. Maestri, A. Mortara, M. T. La Rovere, F. Fanfulla, and P. Sleight, "Periodic breathing in heart failure patients: testing the hypothesis of instability of the chemoreflex loop," *J. Appl. Physiol.*, vol. 89, no. 6, pp. 2147–2157, 2000.

[14] I. Santamaria, P. P. Pokharel, and J. C. Principe, "Generalized correlation function: definition, properties, and application to blind equalization," *IEEE Trans. Signal Proc.*, vol. 54, no. 6, pp. 2187–2197, 2006.

[15] A. Porta, S. Guzzetti, R. Furlan, T. Gnechchi-Ruscone, N. Montano, and A. Malliani, "Complexity and nonlinearity in short-term heart period variability: Comparison of methods based on local nonlinear prediction," *IEEE Trans. Biomed. Eng.*, vol. 54, no. 1, pp. 94–106, 2007.

[16] J. W. Xu and J. C. Principe, "A pitch detector based on a generalized correlation function," *IEEE Trans. Audio, Speech, Language Proc.*, vol. 16, no. 8, pp. 1420–1432, 2008.

[17] R. Benitez, E. Alvarez-Lacalle, B. Echebarria, P. Gomis, M. Vallverdu, and P. Caminal, "Characterization of the nonlinear content of the heart rate dynamics during myocardial ischemia," *Med. Eng. & Physics*, vol. 31, no. 6, pp. 660–667, 2009.

[18] A. Gunduz and J. C. Principe, "Correntropy as a novel measure for nonlinearity tests," *Signal Proc.*, vol. 89, no. 1, pp. 14–23, 2009.

[19] T. Schreiber and A. Schmitz, "Surrogate time series," *Phys. D*, vol. 142, no. 3-4, pp. 346–382, 2000.

[20] A. Garde, L. Sörnmo, R. Jané, and B. Giraldo, "Correntropy-based analysis of respiratory patterns in patients with chronic heart failure," in *Engineering in Medicine and Biology Society, 2009. EMBC 2009. Annual International Conference of the IEEE*, Minneapolis, Minnesota, USA, 2009, pp. 4687–4690.

[21] A. Garde, L. Sörnmo, R. Jané, and B. Giraldo, "Correntropy-based spectral characterization of respiratory patterns in patients with chronic heart failure," *IEEE Trans. Biomed. Eng.*, 2010.